Fingerprints of Colonization
Chloe Shawah

Throughout my first few semesters in college, I kept hearing about the "effect of colonization," and through this project, I wanted to use data science techniques programming in R to quantify and visualize this effect. Thus, the guiding research questions for this project are:

1) Are there differences between the current success/stability of countries that gained their independence recently (countries that were colonized until the last few decades) and those that have been independent for centuries?
2) Have different colonizing powers left their colonies in different conditions?
3) Which colonizing powers left their colonies in the best shape?

To answer these questions, I first found an essential dataset from the Issue Correlates of War Project (ICOW) led by Paul R. Hensel at the University of North Texas titled "Colonial History Data Set." This data set contains the year and month that nations gained their independence, as well as the government/nation they gained their independence from. As Hensel himself explains, there are challenges and subjective judgement calls that arise when trying to make something as complicated as foreign affairs fit into a tidy dataset. Sometimes, it is difficult to determine which foreign power to call the "colonizer" and to mark an exact year when a country became independent. My project defers to all of the judgement calls made in Hensel's dataset, but I encourage viewers to keep in mind the intricacies that are eliminated when forcing a very complicated history into a data frame, and to think critically about which countries have been listed under which colonizing powers. Further, as I began titling the visualizations in my project, I changed the label "colonizers" to the more general "foreign occupiers," as it did not seem appropriate in all cases to call the occupying country a "colonizer."

After finding this dataset to form the backbone of my analysis, I needed to find a way to compare the success or stability of countries. Once again, measuring the "success" of a country is far more nuanced than a single number, so I opted to look at "success" in a multi-faceted way, looking at economic prosperity, health, education, and overall quality of life separately. For each of these categories, I found an indicator to use as a metric:

- Economics: GDP per Capita in PPP (scaled as log-base-10 in visuals), obtained from the World Bank.
- Health: Child mortality rate, deaths under 5 y/o per 1000 live births, obtained from the World Bank
- Education: Literacy rate, obtained from the World Bank
- Overall Quality of Life: Human development index, an aggregate metric of many indicators of quality of life, obtained from the United Nations Development Programme

For my first visualizations, I plotted the value of each indicator in the past year against the year the country gained its independence and colored the points, each representing one

country, by their last foreign occupier. The scatter plot for the GDP per capita indicator is included below:

GDP per capita in 2018 by Year of Independence/Formation
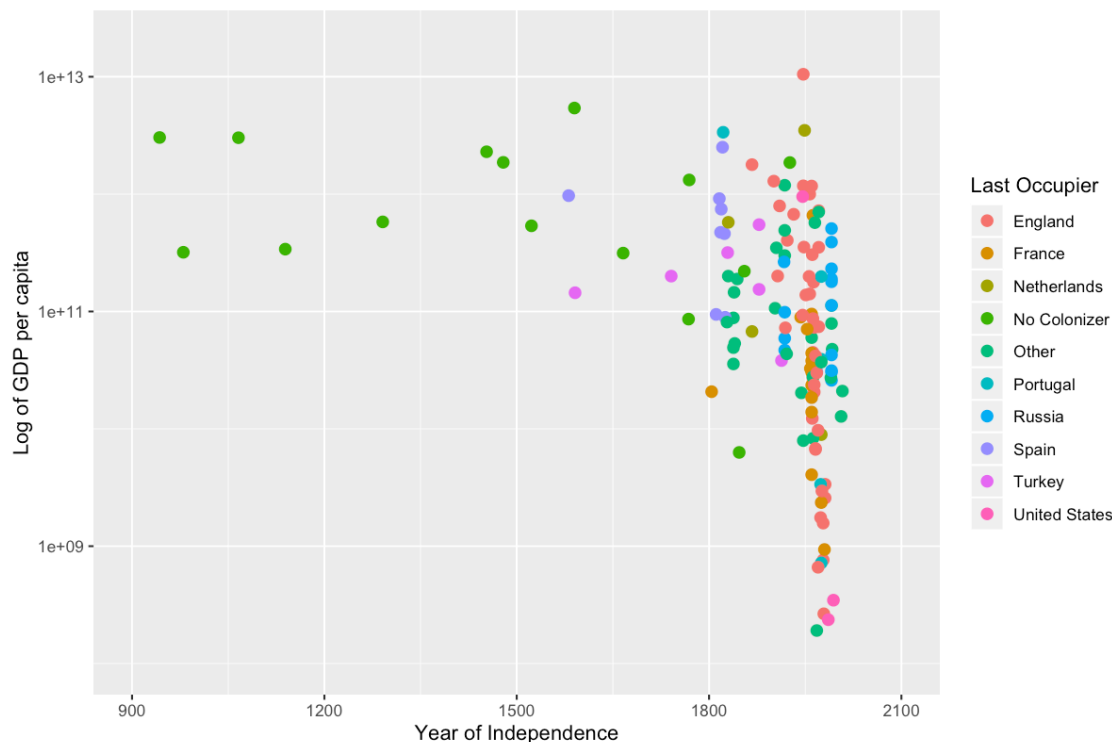

We can see from this graph that there is an interesting correlation between independence year and current economic prosperity. If a country was independent before 1900, its GDP per capita is fairly high. For countries that gained their independence later, however, the current economic state is far more irregular, and varies from very low to very high. Perhaps this indicates that countries needs time to recover from occupation by a foreign power, though we cannot be certain with this graph alone. We can also see by the color of the dots that there does appear to be some differences in outcome based on last occupier.

I repeated this scatterplot with the other three indicators and observed similar trends. It is important to keep in mind when viewing the other graphs that high GDP is good, high literacy rate is good, and high human development index is good, but high child mortality rate is bad, so this child mortality graph appears to be reflected.

To investigate deeper, rather than looking at just a snapshot of todaty I plotted the GDP per capita over time of world nations and faceted these lines into plots by last occupying power. These graphs are included in my project's ShinyApp to allow viewers to view trends over time for individual countries and see how they compare, but I will not include those visualizations here for the sake of space.

Then, for each year of data logging, I averaged the GDP per capita of all the countries that had been occupied by a given occupying power, which yielded an average trend line of average GDP per capita over time of countries last occupied by country "X." This then allowed me to create the ultimate report card for the forgeign occupying powers by plotting these average trend lines against each other for each indicator. This process was created for the child mortality and human development index indicators as well, but had to be left out for the literacy rate indicator, as there was not enough data collection to create a coherent visualization. This plot of average trend lines for GDP per capita is included below:

Tracking Average GDP per capita over time of Counties by Last Occupier
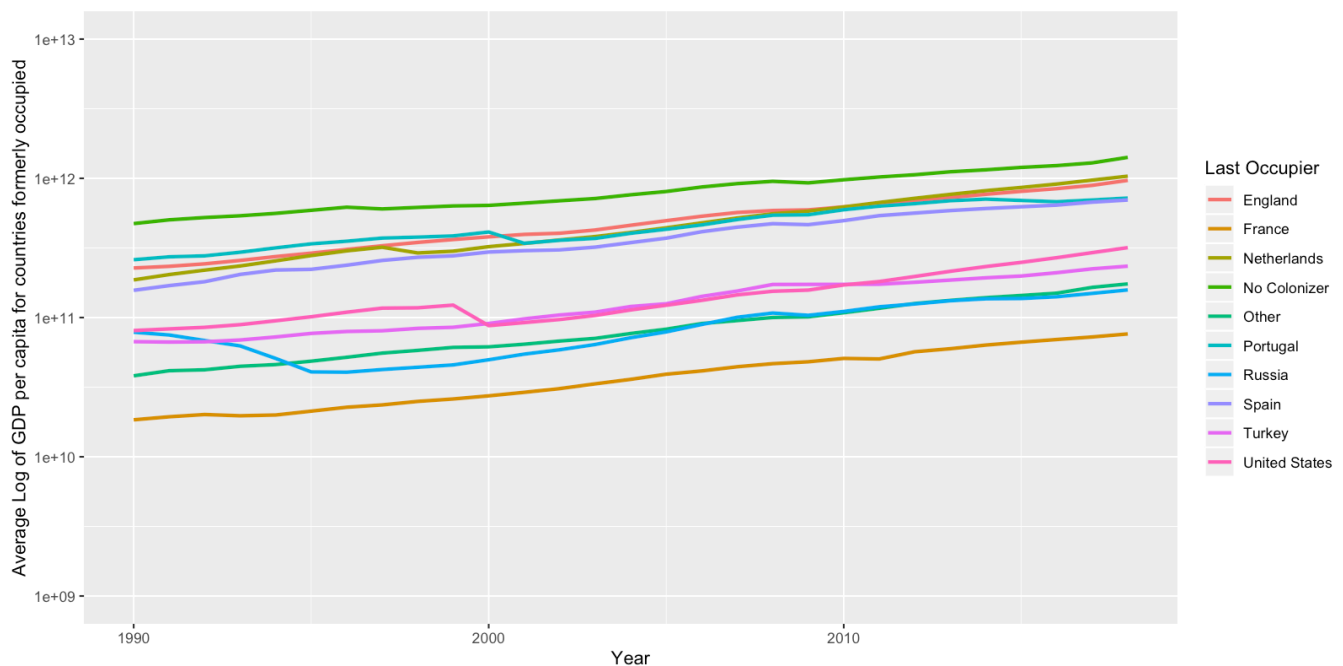
This graph demonstrates the infamous "effect of colonization" in a fairly simple way, just by taking the average GDP per capita of countries last occupied by different occupiers over time. Even with just this simple method, we can see that a gap does exist, and is actually more pronounced than I was expecting. The green line is the average GDP of countries that had never been colonized, and we can see that this average has been consistently higher than the averages of countries that have been colonized. The brown line at the very bottom represents the countries that were last occupied by France, perhaps suggesting that France left the countries it colonized in the worst shape.

Once again, this graph is repeated for the health and quality of life categories (not education due to sparse literacy rate data), and shows variations of a similar trend, with the last occupiers ranked in slightly different orders.

Thank you for taking the time to read about and look at my project. I hope it has provided some explanation and background for a concept, the "effect of colonization," that can at times feel a little abstract.